Math 281A Homework 4 Solution

1. Let $\{X_i\}_{i=1}^n$ be an i.i.d. sample from Poisson distribution with mean θ. Find a variance stabilizing transformation for the sample mean and construct a confidence interval for θ based on this.

   Solution: By central limit theorem,

   $$\sqrt{n}(\bar{X} - \theta) \xrightarrow{d} N(0, \theta).$$

   The transformation

   $$\phi(\theta) = \int \frac{1}{\sqrt{u}} \mathrm{d}u = 2\sqrt{\theta}$$

   is variance stabilizing. So, a confidence interval with level $1 - 2\alpha$ for $2\sqrt{\theta}$ is

   $$\left(2\sqrt{\bar{X}} - \frac{\xi_\alpha}{\sqrt{n}}, 2\sqrt{\bar{X}} + \frac{\xi_\alpha}{\sqrt{n}}\right),$$

   and a corresponding confidence interval for θ is

   $$\left(\left(\sqrt{\bar{X}} - \frac{\xi_\alpha}{2\sqrt{n}}\right)^2, \left(\sqrt{\bar{X}} + \frac{\xi_\alpha}{2\sqrt{n}}\right)^2\right).$$

2. Let $X_1 \sim \text{Uniform}(0, 2\pi)$, and let $X_2 \sim \exp(1)$, independent of X_1. Find the joint distribution of $(Y_1, Y_2) = (\sqrt{2X_2}\cos X_1, \sqrt{2X_2}\sin X_1)$.

   Solution: The joint density of (X_1, X_2) is

   $$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi} e^{-x_2}.$$

   The determinant of Jacobian matrix for $(Y_1, Y_2) = (\sqrt{2X_2}\cos X_1, \sqrt{2X_2}\sin X_1)$ is -1, and $x_2 = (y_1^2 + y_2^2)/2$, so

   $$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}} |\det J| = \frac{1}{2\pi} e^{-\frac{y_1^2+y_2^2}{2}}.$$

3. Let $\{X_i\}_{i=1}^n$ be i.i.d. from logistic distribution with cdf

   $$F_\theta(x) = \frac{e^{t/\theta}}{1 + e^{t/\theta}}, \quad \text{for} \quad t \in \mathbb{R}.$$

   (a) Find the asymptotic distribution of $X_{(n)} - X_{(n-1)}$;

   Solution: Since $F_\theta(x)$ is continuous and invertible with

   $$F_\theta^{-1}(u) = -\theta \log(1 - u) + \theta \log u,$$

   we have

   $$\begin{bmatrix} -\theta \log(1 - U_{(n-1)}) + \theta \log U_{(n-1)} \\ -\theta \log(1 - U_{(n)}) + \theta \log U_{(n)} \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} X_{(n-1)} \\ X_{(n)} \end{bmatrix}, \tag{1}$$

   where $\{U_i\}_{i=1}^n$ are i.i.d. from $\text{Uniform}(0, 1)$, and $\stackrel{d}{=}$ stands for equivalent in distribution. Furthermore, it's not hard to check that

   $$\log U_{(n-1)} \xrightarrow{P} 0, \quad \text{and} \quad \log U_{(n)} \xrightarrow{P} 0. \tag{2}$$

Combining (1) and (2) with slight modification gives

$$\begin{bmatrix} X_{(n-1)} - \theta \log n \\ X_{(n)} - \theta \log n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} -\theta \log(Y_1 + Y_2) \\ -\theta \log Y_1 \end{bmatrix},$$

where Y_1, Y_2 are i.i.d. from $\exp(1)$, and the convergence comes from an example introduced in lecture. This result leads us to

$$\frac{X_{(n)} - X_{(n-1)}}{\theta} \xrightarrow{d} \log \frac{Y_1 + Y_2}{Y_1}.$$

Then we show that

$$\log \frac{Y_1 + Y_2}{Y_1} \sim \exp(1).$$

This can be calculated directly, since for any $u > 0$,

$$\begin{aligned} \mathbb{P}\left(\log \frac{Y_1 + Y_2}{Y_1} > u \right) &= \int_0^\infty \mathbb{P}(Y_2 > Y_1(e^u - 1) | Y_1 = y) f_{Y_1}(y) \mathrm{d}y \\ &= \int_0^\infty e^{-y(e^u - 1)} e^{-y} \mathrm{d}y = e^{-u}. \end{aligned}$$

Therefore, the asymptotic distribution of $(X_{(n)} - X_{(n-1)})/\theta$ is $\exp(1)$.

(b) Based on part (a), construct a 95% confidence interval for θ. You can use the fact that 0.025 and 0.975 quantiles of standard exponential distribution are 0.0253 and 3.6889;

Solution: A 95% confidence interval for θ would be

$$\left(\frac{X_{(n)} - X_{(n-1)}}{3.6889}, \frac{X_{(n)} - X_{(n-1)}}{0.0253} \right).$$

(c) Simulate 1000 samples of size $n = 40$ and $\theta = 2$. How many confidence intervals contain θ?

Solution: 945 confidence intervals contain θ. R codes are attached:

```
rm(list = ls())
n = 40
M = 1000
theta = 2
cover = 0
set.seed(2019)
for (i in 1:M) {
  X = rlogis(n, 0, theta)
  sortX = sort(X)
  left = (sortX[n] - sortX[n - 1]) / 3.6889
  right = (sortX[n] - sortX[n - 1]) / 0.0253
  cover = cover + (theta >= left && theta <= right)
}
cover / M
```